DOCUMENT RESUME

ED 141 113 SE 022 611

AUTHOR Tripp, D. H.
TITLE The Analysis of Discourse as Evaluation of Productive Thinking.
PUB DATE 77
NOTE 33p.; Not available in hard copy due to marginal legibility of original document

EDRS PRICE MF-$0.83 Plus Postage. HC Not Available from EDRS.
DESCRIPTORS Creativity; Creativity Tests; *Educational Research; Elementary Secondary Education; *Evaluation; *Interaction Process Analysis; *Mathematics Education; Problem Solving; *Productive Thinking; Test Construction; Tests

ABSTRACT

This paper provides a thorough description of a method of analyzing and scoring group discussions from a particular point of view. After discussing shortcomings of traditional methods of reporting data from group discussions and problems inherent in the use of paper-and-pencil creativity tests, the author describes a method which was developed as a part of the CORT Thinking Project (Cambridge, England). The method calls for the listing of points made in discussion as both rows and columns in a matrix, and using the cells of the matrix to record relationships among points. Columns tallies are then made for various categories such as example, judgment, digression, etc. After this matrix is completed, a scoring matrix is developed. The construction of the scoring matrix is a difficult step in the evaluation. New scoring matrices must be developed for each discussion topic. (SD)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED141113

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

The Analysis of Discourse as Evaluation of Productive Thinking.

D.H. Tripp
University of Cambridge
Department of Education
1977

SE 022 611

1.1 The method of language interaction analysis described in this paper was developed as a method of comparison of the thinking of two individuals or groups, or of the same individuals or groups upon different occasions. Unless such comparisons are feasible it is impossible to determine the extent to which various changes are taking place as a result of maturation, education or training. The search for a direct method of comparison was the search for a means of converting written or spoken words into indices of thinking ability. That search turned upon the necessity for a means of handling language interaction more manageably than existing methods allowed.

The purpose of this paper is to outline the mechanics of the analysis in such a way that the reader will be able to replicate the method for himself and to illustrate the use of the method with reference to the evaluation of children's thinking behaviour for which it was originally designed. Thus it is outlines work original in two aspects:

(1) the first attempt to produce a codified model of a group discussion in terms of both form and content

(2) the attempt to reduce children's interactive thinking behaviour to a set of comparative indices that is not simply a series of test items.

Problems of analysis of discourse

1.2 Previous research on group discussions have involved an initial verbatim transcript of the discussions, followed by an intuitive search for pattern and meaning expressed in selected quotations and abstracted totals. For example, Barnes (Barnes, Britton and Rosen, 1969) examines teacher questioning strategies, expressing his finding in tables and illustrating the categories of the tables by direct quotation. There are three major disadvantages with this method:

1.2.1 (1) A transcript is lengthy and expensive to produce, both in time and cost. If all the data has to be transcribed, it limits the amount of data used.

1.2.2. (2) Whilst the transcript provides too much information, the findings and totals provide too little.

For instance, in the perception of a spiral pattern, Britton (Barnes, Britton and Rosen, 1969) can only give us selected quotations to illustrate his findings. Both Flanders (1965) and Sinclair and Coulthard (1975) are limited by the use, either of a full verbatim transcript, or of totals abstracted from that transcript. The problem with totals is that it is impossible to check any or all of the individual items that make up that total, and impossible to see how they related to other items in other totals, unless the author has already done it by selected quotation, or having the transcript before us, we are prepared to devote the time to tease out the information for ourselves. Even having done that, there is no better way to communicate our conclusions than the totals or selected quotations available to the author in the first place. Furthermore, as it is impossible to relate all aspects to all other aspects, the investigation of patterns of interaction must largely be predetermined by the particular patterns the researcher or reader had previously decided to investigate. It is difficult to identify new, important or related variables without an overall model of the discussion.

1.2.3 (3) But the greatest disadvantages are a transcript can never be made to model the discussion. A tape is sounds in linear time and a transcript signs in linear space. Either way, aspects are only proximal if they happen to have occured proximally in time, whilst for the perception of the interrelation of aspects one needs to be able to see all the essential features of the whole discussion simultaneously. Thus a model is necessary, and a transcript will never do.

1.2.4 What has been developed is simply a number of conventions to enable the content (what was said) and the form (how it was said) to be coded or "logged" into a matrix. The reason for including form, content and analysis on a single matrix is to enable researcher or reader to simultaneously observe what a particular point is, when and how it is made, how it is analysed and how it relates to the form, content and analysis of any other point in the discussion.

In all the examples included in this paper a ten minute discussion between four to six participants has been coded into a single matrix.

PRIMARY MATRIX

"Should pupils be paid a small wage for going to school?"

Group 3.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1. School is for learning	1						7											
2. Pupils need for money		2	3					8						14				
3. Needs depend upon job or expenses			3	4														
4. ie. travel to training away from home				4														
5. Pupils need money, but less than teachers					5							12	13					
6. To pay would cost too much.						6						12						
7. Pupils learn for themselves without money							7											
9. Teachers would not like and object									9		11							
11. Teachers would strike											11							
12. (5) would be too much												12	13					
13. Even if payment less than £1 per week													13					
14. Pupils do not need money at school														14	15			18
15. Money needed later for marriage															15			
16. Parents could pay instead																16	17	
17. Parents should be consulted																	17	
18. Pupils need money to travel to school																		18
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
+, −, 0.	−	+	+	+	0	−	−	+	−	−	−	−	−	−	−	−	0	+
New Topic / Development	X	X	2	3	X	X	1	2	X		9	6	12	2	14	X	16	14
Repeat / digression																		
Counter							7											14
Specification					X								12					
Example			3												X			X
Judgement	X						7			X								X
OPV		X	1	2			2	2	X		9			2	2		X	2
C&S									X		9	X						
APC																X		
Speaker	1	1	1	1	1	1	1	3	1	2	1	1	1	3	4	2	1	2
Tape Reference	3	8	11	17	20	31	36	38	46	48	58	60	62	74	76	91	93	97
Time (by 15 secs.)	1	3	4	6	9	10	12	13	15	16	19	20	21	24	25	30	31	33

LOGGING THE POINTS

2.1.1 Points are placed in the matrix reading in chronological order of occurence from left to right. This means that point number always equals column number. Content is summarised in the rows, and the rows are numbered according to the number of the point containing that new content. Were a discussion only to contain points using new content, there would be as many rows as points. But in practice people repeat themselves or just make judgements, which means that there are more columns (points made) than rows (new content).

2.1.2 A point employing new content alone will have only one number in the column, and that number will be opposite the row summarising the new content. But often a point will be made referring to, employing or developing content already mentioned in a previous point, in which case the column number will also be placed in the row where that previous content is summarised.

Thus, in Example 1, the first point is logged in column 1 and its content summarised in row 1. Point 2 is independent of point 1 and so refers only to new content which is summarised in row 2. Point 3, however, develops from point 2 and is so logged with reference to point 2, as well as a summary of its new content on row 3. Similarly point 4 refers to rows 3 and 4. Later, point 8 simply repeats point 2, so it is logged only in row 2, there being no new content. The best way of explanation is to illustrate how the first five points were logged by reference to a verbatim transcript:

2.1.3 Point 1 "No, not really, because they've got to go to school to learn and if they're going to get money it's not worth while the teachers teaching them."

Point 2 It would be worth it if you needed the money for things you had to do.

Point 3 That depends on what job you're training for, what expenses you have.

Point 4 Like if you were training to be a teacher in Southampton and you had to get there.

Point 5 You'd need some money, but not so much as the teachers, because the teachers teach you at school so you can learn."

1.4 The number of new input rows is not analysed as such, so when in doubt use a new row. All tapes should be heard at least three times, so a rough first copy can always be amended later, and it is easier to delete too much informaton than to add more. The reason for the input of content rows is to make the matrix intelligible, so one can deduce the content of a point registered in the matrix. The rows should therefore contain all the content of the ideas expressed, including all the variations of the ideas.

2.2 Recording the position of points on the tape.

2.2.1 One of the advantages of the matrix is that it obviates the necessity for a verbatim transcript of the discussion because the content of the point made is summarised. One does, however, sometimes wish to know what was actually said, in which case it is important to be able to quickly find the precise position of the point on the tape. Although it is easier to find this from a transcript, the transcript will only contain the words used; there is much more information on the tape, and the subtleties of intonation, pace or expression are relevant. So, even though a transcript may have been made, the position of each point on the tape must be known. There are two ways to locate it: "Tape Reference" and "Time".

2.2.2 Tape Reference.

The reading of the counter on the recorder is set to zero at the start of the discussion, and the reading at the end of every point logged. It is important that the reading should refer to the end of the point as one usually wants to listen to the whole of the point but not to begin the next. Stopping the machine when the counter registers the end of the point is the most accurate means of ensuring that the correct point is heard. Note that this reading will be different on different machines.

2.2.3 Time.

Because the counters are different on different machines, even of the same make and model, a second reference is made. Having completed the matrix and analysis by listening to the tape at least twice, listen to it a further time with a stopwatch to record the quarter minute in which each point was completed.

This not only gives an indication of the position of the point, but the length of time for expression and the chronological frequency of points at any stage in the discussion. (see graph 1) Furthermore, it allows the position of a point to be found on a different machine simply by leaving the tape to play to the required quarter minute.

2.3 Speaker.

2.3.1 With a good stereo recording it is usually possible to tell which point is made by which speaker. Where the discussion is fairly evenly distributed between the participants it is more difficult, but then, unless one is analysing every individual's contribution, it is less important to know which points were made by whom. In the example Speaker 1 dominated the discussion, under which circumstances speaker identification was particularly easy.

2.3.2 If one is interested in the comparison of individuals' participation then it is appropriate to compare their contribution in different discussions, as in a pretest/posttest comparison, for instance. Not only should one look at their contribution in terms of the number of utterances, but also in terms of the other classifications. Do they always disagree with another participant? Do they bring new ideas to the discussion? What is the sum score for their ideas? Some comment by the researcher is also necessary to inform the interpretation of these figures. Does the dominant speaker interrupt others, or does he speak in the absence of any other contribution? Interruptions could be scored on another row, or the speaker's number ringed. Periods of silence may also be given columns and their duration scored. The contribution of individuals may be examined in this way as well as whole groups. It is useful to be able to see the part played by a particular pupil in a discussion when, for instance, the teacher thinks, from observation in the lesson, that the pupil has changed in his thinking or group behaviour.

2.4 The above aspects are basic to any other factors that are to be analysed from the tape. This paper simply demonstrates a particular use of the method in the analysis of thinking, but it is equally suited to the analysis of other factors, in other situations, such as teacher/pupil interaction variables analysed by Flanders (1965) and Sinclair and Coulthard (1975). In the

latter method, for instance, all one would do would be to place the comment summary in the rows, record the time in shorter units, the speaker and the point as above but, instead of the "thinking" categories outlined in this paper, use their categories of "Classes of Exchange" and "Classes of Acts".

3 The analysis of thinking

3.1 This analysis was developed specifically to measure the effect of the CoRT Thinking Project upon children's thinking. The content and method of the project has been well documented elsewhere (de Bono, 1973, 1975. Tripp, 1976) and although it is not within the scope of this paper to discuss the background, some indication of the nature of the training given is necessary to interpret the analysis. The author of the project, Edward de Bono contends that, if the majority of teachers feel that pupils should learn to think through learning the subjects of the curriculum, then they may be able to learn to think much better if the skills and operations involved are identified and taught as such, instead of being left as a vague and hoped for outcome of teaching other things. The project is designed as a structured framework of thinking skills that are taught separately and directly. The course is defined solely in terms of these thinking skills which are practised on information chosen simply to illustrate the application of the skill. In all the lessons the method is to explain the skill of process, illustrate it in action, and to practise using it in small group discussions. All the thinking training is therefore corporate.

3.2 As evaluator, I was then faced with a very complex and fluid situation where thinking was expressed in verbal interaction and where the use of a number of different processes would only be likely to appear implicitly in the way ideas were generated. At the beginning of the evaluation in 1972 there was no valid way of measuring changes in this kind of situation.

3.3.1 Some teachers were tempted to use standard "creativity" tests, such as the Use of Objects Test popularised by Liam Huudson (1968) In this test the subjects are asked to find as many ways as possible to use an object such as a brick. The major shortcoming of the test is that there is particular technique to produce ideas. One simply associates the brick with other objects: "brick/desk", for example immediately gives rise to:

prop for leg; paper weight; conversation piece; book end; drawer divider; etc. As a test of "creativity", discovery of the technique invalidates the test, for it is an indirect measurement. One would hardly say that a person who used the technique to score highly on the test was more "creative" than the person who did not. Although one could use the test to measure a person's ability to use the technique, such a measure would not reflect any change in their creative abilities, because the technique merely enabled them to alter the measuring instrument.

3.3.2 A similar example to illustrate the problem of indirect measurement would be to measure the weight a person could lift by reading the weight from scales. If the person discovered the technique of adjusting the scales so they registered a greater weight than was actually being lifted, one would no longer be measuring their ability to lift weight, but their ability to alter the scales. We were concerned to find a way to measure thinking directly so that any improvement in technique would be an improvement in thinking ability. To continue the weight lifting example, we wanted to measure the weight a person would lift by looking at the weight lifted, not at the scales. We also wanted to look at the technique employed to lift it, so that to "teach a test" (i.e. how to lift more weight) would not invalidate the test because more weight would actually have been lifted.

3.4 To be able to evaluate a person's thinking performance rather than this test performance is not only important because techniques to improve performance are employed, it is important to establish a means of looking at any thinking situation because it is hypothesised that thinking performance varies according to the situation, and any indirect test can only measure performance on the particular test.+ The development of "tests" was seen as the development of a method of analysing examples of thinking; the "test situation" was seen as no more than the

---

+Some readers may already be familiar with the "J - test". I developed this as an exception to the principle of direct measurement, so it is open to all the criticism relating to all indirect measures. It should therefore be stressed that it is not intended as a test of general thinking performance, but a test of a single aspect of a subject's approach to new ideas: their readiness to make or defer judgements. Tripp, D.H. (1976

collection of pupils' written work, or the recording of a discussion. It is not entirely true to say that there is no "test situation" as such, for although writing may be collected without the subjects' knowledge that they are being tested, with group discussions it is difficult because a special situation has to be created in order to obtain a recording of a sufficiently good quality for analysis. However, in experiments where it was sought to establish the effect of the training in thinking skills as opposed to the method of group thinking, control groups were given equal opportunity for group discussion during the intervention period.

3.5.1 The Definition of "Point"

Our early analyses relied upon the "point" as a unit of thought. One group made more "points" than another and those "points" covered more ideas. Yet it was never possible to define quite what a "point" was. One subject said "X" and later "Y". Another subject said "X but Y": should the former be scored as two points, the latter one? What happens when D entailed A, B and C and one subject simply made point D and another subject made points A, B and C? Some "points" could be called "sub-points" of the same basic idea: one subject may ask, "How would others feel about M? but another may ask, "How would X feel about M? How would Y feel about M? ..... It was impossible to define what actually constituted a point and how one point should be scored in relation to the total. The point score was not only largely arbitrary, but suffered from all the problems of evaluation such as those dealt with by Jackson and Messick (1965)in relation to creativity tests.

3.5.2 It is however necessary to be able to divide up a whole discussion or stream of utterance into manageable units for analysis. One cannot score a whole conversation in a single operation. Attempts were made to divide thinking into quite arbitrary sections (such as single sentences and five seconds) without much success, when it became obvious that the "point" was the natural arbitrary division for thought, but should not be the major unit of quantitative comparison.

3.5.3 So present scoring methods rely on the point only as a convenient unit to analyse, not as a score. It does not then matter if a subject has made two points of an idea that another subject has made into one. If it is simple, analyse it as one; if complex,

different numbers of points according to the amount of information they can analyse in one unit. Providing the sum of the units is the same it makes no difference: what one is scoring is the number of new inputs, the number of developments, the number of people considered, of which there may be none or several in a single point. So, with the example of two subjects who both use the new inputs A and B, one using them together, the other separately, the question as to whether they are one or two points is irrelevant: A and B are scored as exactly the same inputs in both cases, and further, that in one case they were used together and the other separately is also shown by the matrix.

3.6 THE X AND N CONVENTION

In the analysis of points it will be noticed that points are sometimes logged with a cross (X), sometimes with a number (N). The convention is that the X indicates that the element of the point refers either to the point being scored, or to the question under discussion. The N always refers to another point, and is the number of that other point. Thus, in the "Example" row, an X means that the point contains an example of what the point is about: "I think children would waste the money, they'd just buy sweets." The new topic is given with an example. Were "I think children would just waste the money." an input from speaker D, and "Yes, they'd just buy sweets" a development from speaker E, then they would be logged as separate points, and E's contribution be analysed by the number of D's point in the Example row, for E's point would be an example of D. In the example analysis point 4 is an example of point 3, and so is logged with 3's N.

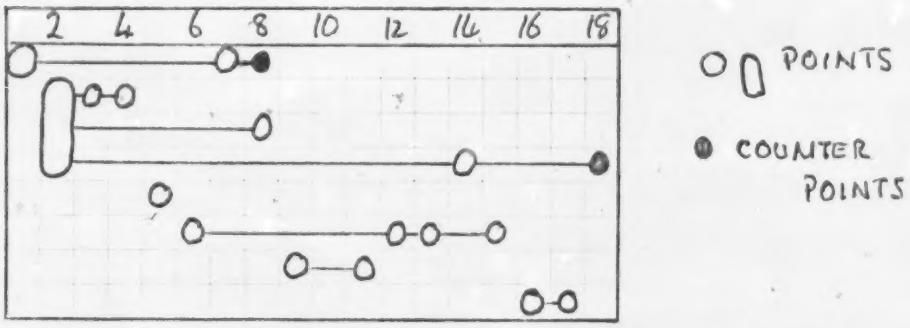
3.7.1 New Topics and Developments

The X and N convention applies to this row. The importance of the row is that it is an indication of how the points raised are related, both logically and in context. Seldom do participants mention the point that they are developing or assuming, but the antecedents of a point are usually obvious in the content analysis. The distinction between the New Topic and a Development is that the former has no antecedents in the discussion, whereas the Development is logically or contextually related to a point already made. In the example, note points 3 and 4. It is not difficult to see that 3 is a development

of point 2, 4 a development of point 3.

3.7.2 Various factors have to be considered when making these relations. Sometimes there is no clear logical connection in the content, though the context shows connections. For example, in the experiment below, point 38 in group 3 connects points 28 and 37, which appear "logically" unrelated but in fact were related by the formation of a "A but B" type point where "A" is "Would help poor parents" and "B" is "But their rates would also increase". Similarly a point may be developed as an interruption of an unfinished point, where the completed point bears no apparent relation to the point made as an interruption. It is important to take these "contextual" clues into account, and to express them in as much detail as possible.

3.7.3 These relations may best be expressed visually in a flow diagram of points.



3.7.4 Plotting the New Topics and Developments in this way gives an indication of "key" ideas. That is, those ideas that were particularly fruitful in terms of the number of other ideas that were developed from them, or particularly important in that they were constantly referred back to as the discussion developed.

3.7.5 Note that this information could not easily be extracted from the Primary Matrix. Look at point 2, for example: on the Primary Matrix the cluster is not obvious; point 4 is only connected to point 2 through its connection to point 3. Yet in the actual discussion (see 2.1.3.) they are directly derived from each other. This is quite clear in the diagram above.

3.7.6 Differences in style of discussion can also be revealed by the diagram. In this example the proportion of points derived from the key idea in relation to the total number of point is about a third. But a group generating many new topics will contrast sharply with a group developing a few new topics.

3.7.7 Similarly, the perception of the "spiral" nature of the development of ideas (Britton, 1969) is clearly shown on this diagram. In the above example all the major ideas (beginning with points 1, 2 and 6) developed in this fashion, and, apart from the fact that one can see this very clearly, one can see how each of the developing ideas related to the rest of the discussion.

3.8.1 REPEATS AND DIGRESSIONS

These obey the X and N convention, the X can only refer to a digression the duration of which is recorded; and the N must always refer to a repeat, as it will be the number of the point that is being repeated. Should a digression be repeated it is scored as an X, partly to distinguish it from a repeat, but also so the content of the digression is shown by the column N in the appropriate row. Digressions are quantified by the total number of quarter minutes engaged by them.

3.8.2 Digression and repeats need criteria for definition. A digression is when the discussion veers so widely from the subject that it appears another subject is being discussed. Very often this is an anecdote triggered by an incidental connection. One child may say "If pupils had to take wages home they might be robbed." Another may say, "I know someone who was robbed. One day..." Only the positively identifiable digressions register. If there is general chat around a point which does not establish a new topic in its own right, it is simply ignored. Thus the chat around point 4 in the example does not include any new inputs, nor does it develop the point and it does not establish a new topic, so it is ignored for logging purposes.

3.8.3 Repeats are recorded when a previously made point is simply repeated without effective change for want of anything new to say. This does not occur nearly as frequently as one might expect. Do not record repeats under the following circumstances:

1. When a point was not discussed when it was originally made.
2. When it is repeated as a summary of subsequent discussion ("So prisoners would never go back at the week-end.")
3. When it is used to counter a new statement. (i.e. "You're forgetting that...")
4. When it is repeated as a preamble to a new idea dependant upon it.
5. When it is used as a means of making sense of an idea, a child putting it in his own words.

A repeat is recorded however, if the substance of the original point is made again although the wording is very different.

3.8.5 Although comparisons may be made between groups in the number of repeats or the time spent in digression, these are reflected in the Aggregated Total Score because they carry zero scores and must therefore depress the total score.

3.9 COUNTER

Whenever a point is made against a previously made point it is a "counter". Point 8 repeats Point 2 to counter point 7 which developed Point 1. But Point 2 did not counter Point 1 in that it was not made with the intent of arguing against it. A "countering" point should explicitly contradict an argument, not just put forward another aspect which may or may not be for a different case. The X and N convention applies to this row, the X meaning that a point was made containing an assertion and its counter: "I think it would be a good idea, but my parents wouldn't." would record an X.

3.10 SPECIFICATION

3.10.1 When a point ascribes a particular value to a factor it registers as specific. "We should be paid £1 a day." (amount). Most often a specification reflects the development of or argument over an idea. One person may say, "We don't really need more money." Another may reply, "Yes, we do, (counter) we need £1 a week for bus to school." In that case the factor (needs) was made specific by a consideration of the actual cost.

3.10.2 This category is the residue of some earlier work on the hypothesis that a wider exploration would be expressed in general rather than specific terms. Whilst this does not appear to be the case the category is still useful as it indicates debate about specific details rather than a general exploration of the subject. When a group begins arguing about the exact amount of money they should receive, for instance, they will then have a number of specifications logged. The X and N convention applies, the X showing that a point was expressed with a particular value, the N that it specified the value of a previously mentioned factor.

3.11 EXAMPLE AND INSTANCE

This category is similar to specification, but it registers use of anecdotes and "concrete" thinking. Very often a general point will be made ("Some people need money to get to school") and it will be qualified or expanded by an example or instance ("You know Jacky? She has to come from Mile End, and that costs her 42p a day.") That example would register as both an Example and a Specification: an example of someone who needs money, specifying how much money they need. In practice many N convention Examples are also Specifications.

3.12 Assumptions

Groups may make assumptions about how much the pay for going to school would be: "If we all paid we'd be richer than the teachers." Assumptions are most often made about the structure for implementing an idea: "If they only go to prison at the Week-end they'd never complete their sentence." Assumptions must not be confused with "Consequences". The category is for "logical" assumptions in the generation of the idea, not opinions and prejudices. "Pupils would waste it" is not a "logical" assumption. The consequence "We'd be richer than teachers" is a "logical" assumption because it depends upon particular conditions (large amount) applying to the factor (pay). Similarly, the second example assumes week-end prison sentences would not be finished, which would only be the case if the factor "sentence length in days" remained the same as full-time prison sentences. An "assumption" then is when the idea depends upon particular conditions applying to a factor when those conditions have in no way been considered. It is interesting that the question on pay for school does not generate many assumptions whilst some other questions do.

3.13 JUDGEMENT

Any overtly stated judgement for, against or ambivalent to the idea scores in this row. The first point begins, "No", although modified by "not really", it is still recorded as a Judgement. Sometimes judgements are made in isolation (point 10) in which case they occupy a column. In every case the + - 0 row will indicate the direction of the judgements. The usual X or N convention applies to this row: and X indicates judgement of the idea being discussed; N indicates the number of the point being judged (i,e. point 8 begins, "No, that's silly..." referring to point 7.)

3.14.1 Whilst the above categories are simply features of discussion the following categories are directly derived from the skill training given to the experimental groups by teachers of the project. In other words, they define changes in thinking ability in terms of the project's definition of thinking skills. Although other researchers will have other skill categories and processes that they will wish to monitor, I have only used the project's training skills because I am concerned to find out the effect of training in these skills, rather than to establish whether or not these particular skills represent the basic components of thinking behaviour. So what skills one looks for is less important a consideration at this stage than how one looks for them.

There are five basic skills that I monitor in the examples included in this paper. These are all taught in the first unit of the project under code names.

3.14.2 1. PMI (Plus, Minus and Interest)
This lesson is concerned with the postponement of judgement. Most people judge an idea and then stop thinking about it, or just defend their judgement of it. The Plus and Minus categories are to (a) accept an initial judgement (thus, "I like it" becomes just one Plus point) and (b) to make one go on to look at both sides of the question.
The "Interest" category is used to provide room for points that are not obviously for or against the idea, or that may be both "Plus and "Minus".

3.14.3 2. The CAF (Consider All Factors) lesson is to extend thinking. Most people stop thinking too soon and by trying to CAF room is created for the generation of more points. It can also be treated as a "What have you left out?" exercise.

3.14.4 3. OPV (Other Point of View)
Again most people will think of factors that concern themselves primarily. The OPV trains children to find how many different view points there may be in a situation, and encourages them to see the idea from another point of view.

3.14.5 4. C & S (Consequence and Sequel)
This exercise develops the process of searching for more consequences of an idea and extending thinking further into the future.

3.14.6 5. APC (Alternatives, Possibilities and Choices)
In any situation there are many different alternatives, possibilities or choices which are largely ignored through the immediate grasping of an obvious or habitual point. This exercise trains children to deliberately look for alternative explanations, courses of action and new possibilities,

4 THE USE OF TECHNIQUES

4.1 When assessing the use made of specific thinking processes by a group for comparison between experimental and control groups, it is possible to record for incidence of the products of the processes, hypothesising that where a group produces more consequences, for example, it may be in some measure due to their having been taught to look for them. Sometimes group subjects do specifically mention a process ("But we've not thought of any good points yet") in which case it should be recorded in the researcher's comments; but it is far more likely and usual that any effect of the training will be unintentional finding or more consequences, not because they are consciously looking for consequences, but because they are aware that ideas to have consequences. This general awareness is seen as the residue of the specific training and practice of the lessons.

4.2 The aim of this analysis is to produce indices that relate directly to observed trends and so need little interpretation. Three composite indices have been constructed to evaluate significant aspects of thinking. These indices are calculated from use of the above skills and are indicative of changes that may take place in these aspects.

4.2.1 1. The Balance of Argument Score takes points made "for", or "against" or "neutral to" the making of a case, and combines these as a function of the total number of points raised. It shows the degree of subjectivity with which the topic is dealt, and is directly affected by use of the P.M.I.

4.2.2 2. The Use of Technique Score takes the selected thinking techniques, including the Balance of Argument and Total Point Scores, and expresses the frequency of their use in relation to the expected frequency derived from controls. It shows the effect of training in thinking techniques.

4.2.3 3. The Total Point Score takes all the factors used and values them favourably weighting the use of new factors over the re-use of previous factors. It combines the quality, quantity and number of factors covered, and so represents the CAF operation.

4.3.1 + - O, THE BALANCE OF ARGUMENT SCORE

This category is used to indicate whether a point is for "+", against "-" or neutral "O" to the topic under discussion. The most obvious "O" points are those which consider how the idea might be implemented, or which put forward alternatives to it. As these categories are assigned by the researcher he must avoid imposing his own judgement of whether a point is for or against the argument: always scoring according to the speaker's use of the point. This is usually quite clear, but should a child make a point to support the topic ("We could spend it all on sweets.") it must be recorded as a "+" however clear it may be to the researcher that, objectively, the point is equally against the topic. Such contradictions are made obvious by the nature of previous and subsequent points, judgements and emphasis. They are indicative of emotional handling of ideas, and should be noted for comment. For further accuracy one may employ three more categories; "(+)" and "(-)" for points which are doubtfully "+" and "-"; and x for a genuine unknown. Score "(=)" and "(-)" as half "+" and "-", and x for missing data. Finally, if a point of the "A but B" type has been logged in a single column, assign both "+/-".

4.3.2 Interpretation and comparison for the balance of the argument in "For" or "Against" terms is not simple. Different questions have different ratios of + and - points, so it is not advisable to compare different questions with different groups, although it is not impossible. More often one wishes to compare the same question with matched control and experimental groups, in which case it matters little what the absolute ratio is: the less imbalanced discussion will still have a better ratio than the more imbalanced. I use imbalanced deliberately as I have yet to find a discussion being anywhere near balanced.

4.3.3 As evaluator I was looking for some empirical means of showing whether a group was exploring the possibilities and implications of the questions, or whether they were producing a case for or against it. It seems reasonable to suppose that if a group is exploring ideas they will produce ideas both for and against the

This was overcome by using: $\frac{\text{Lesser of "for" or "against"}}{\text{Total Judgement points}}$

which made the maximum score for a perfect balance 0.5, and increased imbalance decreased the score. The disadvantages of this method were that groups or individuals with low or very imbalanced scores often had zero for their Lesser of "for" or "against" total; also the maximum was less then one.

The former problem was overcome by the arbitrary addition of one to the minimum "for" or "against" score, so an original zero became one, giving:

$$\frac{\text{Lesser of "for" or "against"} + 1}{\text{Total number of judgement points}}$$

4.3.7 TOTAL POINTS

Three interconnected problems remained: the balance calculation favoured the subject with a small score; the scale was still inverse to the number of points made; and the maximum was still less than one. This last problem could obviously be solved by a straightforwad multiplication. All three were eventually solved together. The problem of balance was that the addition of one to the minimum for or against score provided a perfect balance for an individual with three points in the ratio 2:1, for example, but made scarcely any significant difference to the Balance Score of a subject with 18 points in the same ratio. It was not possible simply to multiply the Balance Score by the total number of judgement points made because that total had already been used to divide the minimum number for or against + 1, and would simply have cancelled out. As it was important that the number of non-judgement points did not affect the Balance Score but were affected by it, it was decided to multiply the product of the Balance Score by the total number of separate points made, both judgement and non-judgement. This method also produced a scale with a minimum score of unity and a practical maximum of 20 and so obviated any further calculation.

4.3.8 The final equation for the balance, then, emerged as:

$$\frac{\text{Lesser of "for or "against" points} + 1}{\text{Total number of judgement points}} \times \text{Total number of separate points}$$

4.3.9 With regard to the working of the formula it is interesting to note that the weighting against imbalance and in favour of number of points generated means that a subject who has only generated points in one of either a "for" or "against" judgement category will invariably have a Balance Score of 1 regardless

of the total separate points he has generated. In other words, although a greater number of points generated tends to increase the score, the weighting against imbalance will decrease the score by an equal amount.

4.4 OPV

The expression of a point of view ("Mother would like it" or "Teachers would go on strike") registers as an X in the column when it is first mentioned, and should be recorded to the left of the matrix as New Input. Thereafter mention of the same point of view is recorded by the N of the column in which it was first registered. Thus one has the total number of different viewpoints, and how often each was used. It is the number of different viewpoints used that are counted, not the number of occasions on which a viewpoint was taken. In this way a group considering the question from different viewpoints scores highly in contrast to a group considering only one viewpoint.

4.5 C & S

Any point that considers what would happen as the result of an idea is counted as a Consequence. The idea "Children would be robbed" is a consequence of the subject of the discussion (Should children be paid for going to school?) and would register according to the convention with an X. When a point is the consequence of a previous point more directly than the subject under discussion it is registered by the N of the point it develops. Thus the point "No one would want to get paid then" following the point above would be registered as a development because it is contingent upon the previous point. The defining criterion here is whether the consequence could be made in the absence of the point it appears to develop: if it cannot then it must be a development. As every point containing a consequence which is not a repeat of a previous point must contain a new consequence, all consequences are counted for scoring, whether they develop a previous point, or relate solely to the question.

4.6 APC

Any idea that suggests an alternative to the question or a previous point registers as an alternative and is recorded in the X and N convention. The point, "Children could all receive the same amount of pocket money instead" or "Pupils could be paid for work in the holidays" are alternatives to the question (Should children be paid for going to school?) and so register an X. As with Consequences, Alternatives to previous points are counted with alternatives to the question discussed.

4.7 Use of Minor Techniques Score

4.7.1 A comparison of the raw totals is misleading as a group with more points overall will have more points registered in these categories than will a group with fewer points even if both groups have made similar use of the techniques. Any combined score used in comparison must be a function of the overall number of points that could be the product of these techniques. A simple sum of the rows as a percentage of the overall number of points is more meaningful, but it does not take into account the different proportions in which these techniques occur in control groups and pre-experimental groups, which vary according to age, ability and the question, as experience with a number of analyses has shown. To take these factors into account I calculate the percentage incidence of each technique as an average for the pretest or the control groups and then weight the post-test and experimental group results by a similar proportion, so that a group with double the experimental group percentage incidence of one of three techniques would score double one third of the Use of Minor Techniques Score.

4.7.2 For instance, if the control group scored C&S 20%, OPV 15% and APC 10% giving a total score of 45%, and the experimental group scored C&S 20%, OPV 15% and APC 20%, although they had scored double the control group in one category they would only have scored 5% in all, whereas for doubling a third of the control group total they should score 60. The way to weight the individual scores is to take the most frequent technique as a base, and increase the value of the other techniques to the same amount, so each technique represents one third of the total. Thus, C&S would remain unchanged at 20, but OPV would become $15 \times \frac{20}{15} = 20$, and APC become $10 \times \frac{20}{10} = 20$. The control group total would then be 60. Increasing the experimental group score by the same proportion would give us $20 + (15 \times \frac{20}{15}) + (20 \times \frac{20}{10}) = 80$, which is $(20 + \frac{1}{3} \times 20 \times 2)$ the proportionate increase in the Use of Technique.

4.7.3 Minor Technique Factors

There are always more C & S than any other Minor Technique which means that the OPV and APC scores have to be increased. By taking the mean incidence of all three techniques from 12 control group discussions of the question a factor has been calculated by which to increase the OPV and APC scores. This factor is 2.62 for the OPV, and 3.17 for the APC.

4.7.4 The Use of Techniques Score

By the same process the Use of Minor Techniques Score is extended to include the PMI and CAF processes calculated in the Balance Score and the number of points raised respectively. This is now done by including the Balance Score and Value of Points raised in the calculation, adding them in proportion in the same way that the individual Minor Techniques are combined. As the Balance Score averages one-tenth of the value of the Minor Techniques Score, and the value of points raised averages one half of the Minor Techniques Score, the conglomerate Use of Techniques Score is the sum of:

1. The Use of Minor Techniques Score.
2. The Balance Score increased ten-fold.
3. The value of points raised doubled.

The sum total of these value is known as The Use of Techniques Score.

4.7.5 The Use of Technique Score is important as it shows differences in the way ideas were generated which may be the result of the training. It must be understood, however, that it is a conglomerate score disguising differences between at least three major factors:

1. How effectively the techniques were learned (or taught).
2. How long such training endures for conscious or subconscious "residual" application.
3. How far that learning is transferred to the test situation.

4.7.6 "Expressed" and "Actual" universe

A problem with this approach that has been suggested is that, although a group may produce more Consequences, for example, they may produce so many more points overall that, because the Use of Techniques Score relies upon the number of points generated by a technique as a percentage of the total number of points, ed the Use of Techniques may actually diminish. Initially I consider this indicated less use of that technique (Consequences), but it may be that for any question there are a limited number of points that may be generated by a technique (i.e. the actual universe). Thus, suppose a question has 10 possible alternatives and Group 1 finds 8 of them while generating 30 points overall, their score for use of the APC technique would be 27%, while if Group 2 should find all 10 alternatives but go on to generate 45 points, their use of the APC technique would fall to 22%, although they had found all the possible alternatives. It may

therefore be necessary to find the actual universe for every technique in every question and calculate the Use of Techniques as a percentage of that, rather than as a percentage of the expressed universe.

"Should pupils be paid a small wage for going to school?"

Group 3.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1. School is for learning	1						7											
2. Pupils need for money		2	3					8						14				
3. Needs depend upon job or expenses			3	4														
4. ie. travel to training away from home				4														
5. Pupils need money, but less than teachers					5								13					
6. To pay would cost too much.						6												
7. Pupils learn for themselves without money							7											
9. Teachers would not like and object									9		11							
11. Teachers would strike											11							
12. (5) would be too much												12	13					
13. Even if payment less than £1 per week													13					
14. Pupils do not need money at school														14	15			18
15. Money needed later for marriage															15			
16. Parents could pay instead																16	17	
17. Parents should be consulted																	17	
18. Pupils need money to travel to school																		18

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
+, −, 0.	−	+	+	+	0	−	−	+	−	−	−	−	−	−	−	−	0	+
New Topic / Development	X	X	2	3	X	X	1	2	X		9	6	12	2	14	X	16	14
Repeat / digression																		
Counter							7											14
Specification					X													
Example			3										12		X			X
Judgement	X						7			X								X
OPV		X	1	2			2	2	X		9			2	2			X 2
C&S									X		9	X						
APC																X		
Speaker	1	1	1	1	1	1	1	3	1	2	1	1	1	3	4	2	1	2
Tape Reference	3	8	11	17	20	31	35	38	46	48	58	60	62	74	76	81	93	97
Time (by 15 secs.)	1	3	4	6	9	10	12	13	15	16	19	20	21	24	25	30	31	33

SCORING MATRIX

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
People																		
Pupils		4	0	0			0	0						0	0		1	0
Parents																	4	
Teachers								4			0							
School																		
Government																		
Uses																		
Spend																		
Save																		
Waste																		
Consequences																		
School numbers																		
Work post-school																		
Facilities / Resources																		
Problems								4			2							
Structure of pay																		
Source																4		
Cost,						4						2						
Who paid																		
What pay				4									2					
When Paid																		
Reasons for pay																		
Incentive / Discipline							4											
Others paid																		
School as work	4																	
Needs		4	3	2			0						2	3			2	
Education																		
Alternatives																		
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
POINT SCORES	4	8	3	2	4	4	4		8		2	2	2	2	3	4	4	2

5 THE SCORING MATRIX

5.1 A separate matrix is required to evaluate the value of the ideas used, in the discussion in order to log the content of the points against some standardised abstract of all the relevant ideas that could have been used. This is not possible on the Primary Matrix as it is simply a record of what was said in chronological order. In order to score the multitude of possible ideas some generalisation and amalgamation is necessary. Were this to be the form of the Primary Matrix there would be no record of what was actually said. In practice it is worthwhile making a fair copy of the analysis of the Primary Matrix under the point scores on the Scoring Matrix, as all the information is then contained on the single matrix, but specific details about content can always be looked up on the Primary Matrix. If the two are amalgamated it is helpful to use a different colour for the referencing of the point columns to that of the point values, so the two are not confused.

The Scoring Matrix should mirror the point sequence of the Primary Matrix. Each column on the Scoring Matrix takes the same N and same content as the column of the Primary Matrix, so that column 17 on the Primary Matrix is scored in column 17 of the Scoring Matrix. The only difference is that instead of the content being entered on the left of the matrix in the chronological order of production by the particular group, it is generalised into a standard format, and values placed upon the content of the discussion.

5.2.1 <u>The Construction of the Scoring Matrix</u>

The construction of the Scoring Matrix is a long and tedius task, and every question needs its own matrix. Matrices for some other questions are available and more are in the process of construction.

5.2.2 To construct the Matrix a pool of all the points made by all the groups tested (a minimum of 10 groups is recommended) and these points sorted under general headings such as those heading each section: "People", "Use" and "Structure". Under these general headings are then listed the specific factors that the groups produce in their discussions. For instance, the different "people" whose points of view were mentioned were: "pupils", "parents", "teachers", "the school", "the government" and one or

(5.2.2) two others. It is important to ensure a balance between such detail that there are too many factors used so infrequently that they are not worthwhile, and too few so that genuine new input cannot be distinguished or scored. Thus, all the named categories of people were fairly commonly used, but "people" such as "students", "cripples" and "shopkeepers" were only mentioned once, and so they were amalgamated under "People". Similarly there were several "problems" that were mentioned as consequences, none of which occured frequently enough to merit its own classification. "Problems" included, "strikes", "vandalism", and "robbery". It is important to leave these open categories, and include an "unclassified" category at the bottom for any idiosyncratic factor that cannot be scored in the usual scheme.

5.3 Scoring the Analysis

5.3.1 Initial values of 4 have been put upon these factors. In order to distinguish between a group that is exploring the question widely or in depth, and one that is arguing over details, repeats itself, or is limited to a few aspects, the factors are not re-scored when they are used a second time, and developments and variations of factors score 2. The effect of this is to diminish the overall value of exploration in terms of permutations of factors and to augment the overall value of exploration in terms of a broader consideration of new factors, which is the CAF operation. The group that considered five different categories of "people" will score nearly double the group that considers five different permutations or variations of the category "parents" (i.e. one category of "people"), and the group that considers neither the variations of a category nor different categories will not score at all.

5.3.2 It is not difficult to distinguish between a new factor which is a development of a previously mentioned factor or a variation of it, and a re-use of the same factor which has not been changed in any way. For example "Parents wouldn't like it" would score a 4 for the category "parents" on first use. If it were later followed by "Parents would have to pay through the rates" the subsequent use of "parents" unchanged would not score at all. If however the subsequent use were "Poor parents could not afford it" then "poor parents" would score half the original (i.e. 2) because it developed the factor "parents" in some way. The rule then is to score half the original for any change in a previously used factor, none for an unchanged

factor on its re-use. The sum total of the value of the points scored is referred to as the Point Score.

5.3.3 Note that the Primary Matrix registers the context of a point, relating it to previous factors by the use of the point number in the row where the factor was first introduced to the discussion. As the Scoring Matrix does not show context, but should mirror the first matrix, it is helpful to enter a zero score against a factor that is not developed and so does not score, but is nevertheless within the context of the point, and the zero entry makes the content of the point more intelligible when reading it back. Thus the points 3 and 4 in the example both register the context "pupils" point of view, and although that use carries no score, the zero entry makes its use obvious, and draws attention to the fact that a decision has been made to zero score that use of the factor.

5.4 Evaluating the factors

Compare the three points made by the following two groups:

5.4.1 Group 1

Point 1: "Where is the money to come from?"
Point 2: "Children who come to school but don't work should be paid less."
Point 3: "Parents might think they can have more children so they can get the money."

Group 2

Point 1: "Pupils would all rob each other to get more money."
Point 2: "Pupils would go on strike for more."
Point 3: "Pupils would be richer than the teachers."

In the first case the group has thought about: Source of money, use of pay as reward, effect on population; in the second case the group has put forward two rather unlikely reactions of the pupils, and an outcome of an equally unrealistic assumption about the amount they would be paid.

5.4.2 In order to evaluate the different values of these factors more or less than the standard mark for a factor is scored. For instance, in the example "Needs" were first mentioned in point 2. The next "Needs" (Points 2 and 3) were for work and for money to continue education after school and away from home. The next "Need" (Point 14) was negatively stated that

pupils did not need money at school, whilst the final need (Point 15) was for money when marrying to buy a house. Two of these needs were felt to be important and so scored 3, which was half the original + 1. Had the "needs" considered been the "need" for entertainment or holidays, they could have been scored as half the original (i,e. variation of original) less 1. So what is felt to be a particularly important point or a particularly trivial one can always score more or less, but not more than one mark either way (making a range of 3), and it is most important to keep a record of factors that have received a bonus or a penalty to be consistent. Bonuses or penalties always show up as odd scores (1, 3, 5) whereas unchanged scores remain even (0, 2, 4).

6 THE INTERPRETATION OF RESULTS

6.1 It is relatively simple to compare the performance of two groups on the basis of the three composite indices, which may be tested for statistical significance. It must always be born in mind that these are conglomerate totals and may thus hide wide variations in component totals. It is therefore important that all the components of the indices are expressed, and any conclusions drawn from the composite indices interpreted in the light of their components. Note that the OPV and APC factors will change slightly for different groups. As more data is collected there will no doubt emerge different constants for different variables such as age of group, teaching situation or length of discussion. Different constants have already emerged for different questions, though the differences are rather smaller than expected.

6.2 Pattern analysis is perhaps more illuminating and two devices seem to display differences particularly well.

6.2.1 1. Plotting the cumulative value of points against the duration of the discussion in quarter minute intervals will show very clearly differences in the rate of generation and the content of the discussion points. (Graph 1.)

Group 3

Balance of Argument Score

Raw Totals: + = 5; - = 11; 0 = 2.

Calculation: $\frac{5+1}{16} \times 18$ = 6.65 (6.65 X 10) = 66.50

Minor Techniques Score

Raw Totals: OPV = 3; C&S = 2; APC = 1.

Calculations:

OPV = $\frac{3}{17}$ = 17.65%	(17.65 X 2.62*)	=	46.24	
C&S = $\frac{2}{17}$ = 11.76%			11.76	= 76.64
APC = $\frac{1}{17}$ = 5.88%	(5.88 X 3.17+)	=	18.64	

Point Score = 58.00 (58.00 X 2) = 116.00

USE OF TECHNIQUES SCORE = 259.14

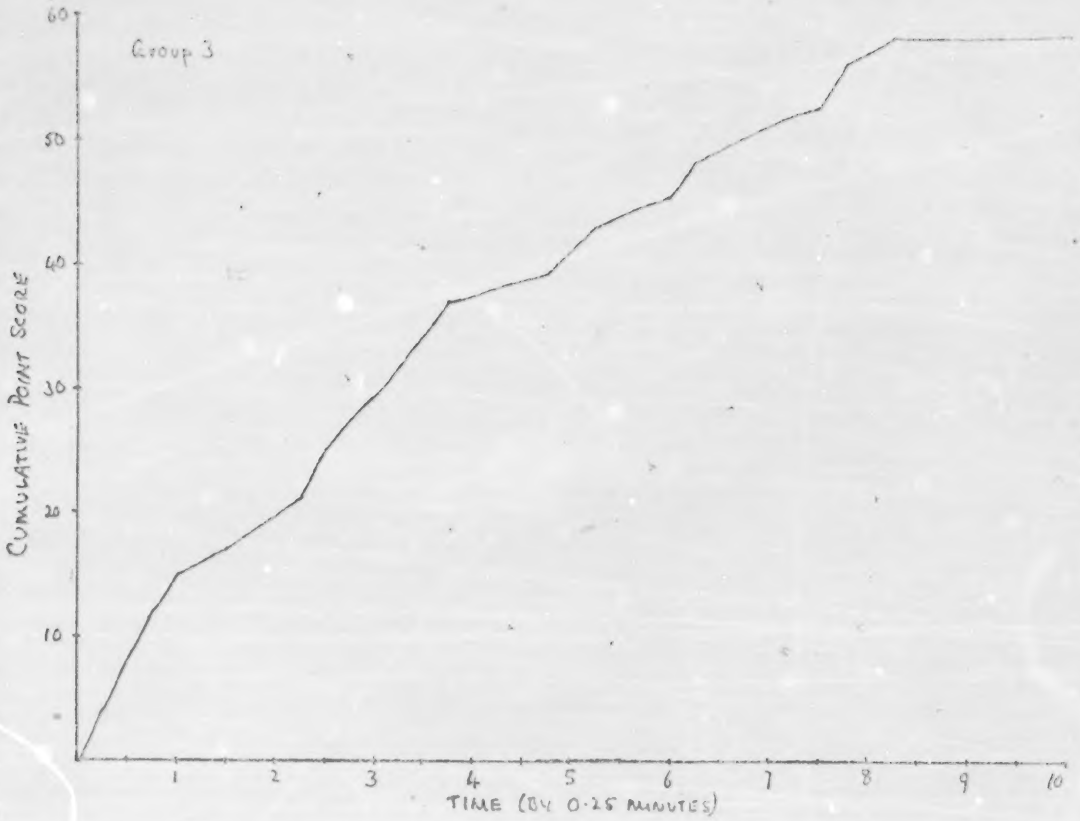
Other Totals

Repeats	1	(5.56%)
Digressions	0	(0.0 %)
Counters	2	(11.11%)
Specifications	2	(11.11%)
Examples	3	(16.67%)
Judgements	4	(22.22%)

* OPV factor
\+ APC factor

NOTES

1. Discussion was terminated by the group with "We've finished." at 8.25 minutes.
2. The mention of "Parents" in point 16 was not scored as it was not bringing in a new point of view, but rather expressed the pupils' point of view.
3. Point 8 was not a repeat of point 2 as it was used to counter point 7
4. Although speaker 1 made 67% of the points he scored 81% of the total point score, an excellent average of nearly 4 per contribution.
5. In so doing he did not prevent others from speaking, but contributed in the absence of any other points. This is shown by the Time, and the slow rate of climb of the cumulative point score on the graph.
6. C&S were very low for this question.



Graph 1.

6.2.1 2. Plotting the connections between ideas illustrates the way in which the points relate to each other (see flow diagram).

6.2.2 It is most important for the researcher to make short comments upon the analysis as it progresses. Although these comments may not tell us much about an individual group, cumulatively they are the best means of discovering new trends for investigation. The "Assumptions" category (10.12) for instance was formulated simply because I was noting points that the groups seemed to accept as valid when I felt they were not. Having noted them I could look back over a number of examples and formulate the category from accumulated data when I had begun with nothing more obvious than a sense that I might be missing something.

6.3 Conclusions

It would not be correct to give the impression that this method of analysis is without difficulties, incongruities and unresolved dilemmas; I have a large file of comments and examples of aspects about which I am far from happy. I have not dealt with

these here as I do not wish to risk turning what I hope is a fairly straightforward account of the method into a discussion of the philosophical ramifications of particular aspects. I hope simply to have given the reader sufficient practical instructions to enable him to analyse his own data.

NOTE

It should not be thought that the work oulined here is the only aspect considered in the evaluation of the CoRT Thinking Project; it is but one aspect of the "hard-data" analysis. An account of the whole evaluation method is to be found in the Evaluation Manual (Tripp, 1977).

REFERENCES

Barnes, D., Britton, J., and Rosen, H., Language, The Learner and the School, Penguin, 1969.

Britton, J. See Barnes, D., above.

de Bono, E., But How Do You Teach Thinking? Times Educational Supplement, 17.8.1973.

de Bono, E., Teaching Thinking, Temple Smith, 1975.

Flanders, N.A., Teacher Influence, Pupils Achievement: Studies in Interaction Analysis, Co-Operative Research Monographs, No. 12, U.S. Government Printing Office.

Hudson, L., Frames of Mind, Menhiun, 1968.

Jackson, P., and Messick, S., The Person, The Product and the Response: Conceptual Problems in the Assessment of Creativity. Journal of Personality Vol. 33, 1965.

Sinclair, J., and Coulthard, R., Towards an Analysis of Discourse, O.U.P. 1975.

Tripp, D.H., Teach Them to Think, Remedial Education, Vol 11, No. 3, 1976.

Tripp, D.H., The J-Test (Deferred Judgement Test), Cognitive Research Trust, 1976.

Tripp, D.H., The Evaluation Manual for Teachers of CoRT Thinking, Cognitive Research Trust, 1976, revised 1977.